

РАЗРАБОТКА АЛГОРИТМА КЛАССИФИКАЦИИ ПРОИЗВОДСТВ ПО ТИПУ ВНУТРИЗАВОДСКОГО КООПЕРИРОВАНИЯ ОСНОВНЫХ И ВСПОМОГАТЕЛЬНЫХ ПРОЦЕССОВ МЕТОДОМ МАШИННОГО ОБУЧЕНИЯ

Задача рациональной организации вспомогательных процессов на предприятии заключается в снижении их себестоимости путем более глубокой интеграции в основной производственный процесс. Цель статьи заключается в разработке алгоритма классификационного анализа для оценки зависимостей между основными и вспомогательными подразделениями и типологии производственных процессов по уровню внутризаводского кооперирования. В качестве метода определения типа производства предложен метод машинного обучения «Случайный лес» с использованием метаалгоритма обучения машин Бэггинга. Разработаны параметры, описывающие затраты на вспомогательные операции, расходы на ремонтное хозяйство и обслуживание оборудования, уровень технической эффективности производства. Апробация алгоритма на примере химических предприятий позволила выделить три типа производств по характеру внутризаводской кооперации процессов по наиболее информативным параметрам. Для оценки полезности и производительности моделей построены диаграммы кумулятивного подъема, где наиболее продуктивным определен тип со средним уровнем внутризаводского кооперирования. Результаты являются первичной диагностикой организации вспомогательного хозяйства, принятия решений о проведении реинжиниринга процессов с целью усиления внутризаводского кооперирования и снижения уровня затрат.

Ключевые слова: вспомогательное производство, производственные процессы, внутризаводское кооперирование, алгоритм, классификационный анализ, машинное обучение, «Случайный лес».

Введение. Предметом исследования организации производства является организация технического обслуживания основного производственного процесса, осуществляемого силами вспомогательных цехов и участков. От качества проектирования производственной системы во многом зависит ритмичность и бесперебойность технологического процесса, техническая и экономическая эффективность производства продукции. Простой основного производства по причине низкого уровня организации вспомогательных производств приводят к снижению выпуска продукции, перерасходу материальных ресурсов, увеличению себестоимости конечной продукции [1, 2].

В практике проектирования промышленных производств используется термин «кооперирование», характеризующий форму организации производства и предполагающий производственные связи цехов, участков, совместно участвующих в производстве продукции [3, 4]. Внутризаводское кооперирование проявляется в обслуживании основных производственных подразделений вспомо-

гательными. Соответственно, чем выше уровень кооперирования основных и вспомогательных процессов и операций, тем больше степень ритмичности и бесперебойности технологического процесса, рациональнее используются ресурсы производства.

Типологию организации производственных процессов по уровню внутризаводского кооперирования предлагается осуществить на основе классификации объектов по ряду признаков. Адекватным инструментом классификации может быть один из алгоритмов искусственного интеллекта. Достаточно широкое применение в решении прикладных задач классификации находят методы деревьев решений — алгоритмы обучения классификатора «Случайный лес» (Random forest), обладающие гибкостью для решения практически любых проблем в области машинного обучения: классификация, регрессия, поиск выбросов и аномалий [5–7]. В сфере организации производства алгоритмы случайного леса используются для принятия решений по модернизации производственных процессов [8, 9], организации контроллинга и повышению ре-

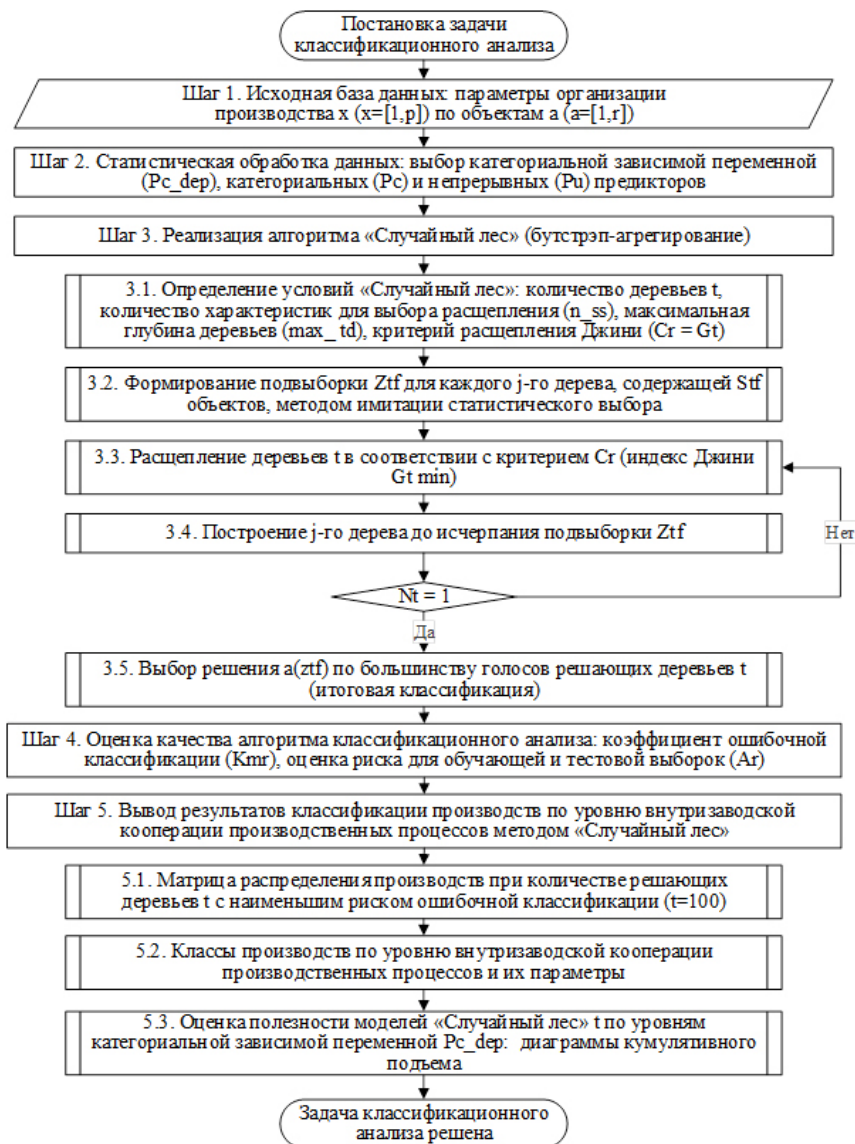


Рис. 1. Алгоритм авторской методики решения задачи классификации производств по уровню внутризаводского кооперирования производственных процессов методом машинного обучения «Случайный лес»

сурсоэффективности промышленных предприятий [10, 11].

Вместе с тем при наличии теоретико-практического материала остаются неизученными отдельные вопросы использования алгоритмов машинного обучения для решения задач моделирования производственных процессов, в частности вопросов организации вспомогательного хозяйства.

Цель исследования. Цель статьи заключается в разработке алгоритма классификационного анализа для оценки зависимостей между основными и вспомогательными подразделениями и типологии производственных процессов по уровню внутризаводского кооперирования.

Методика исследования. Разработана методика и выстроен алгоритм действий, включающий последовательность математических и логических процедур выбора производств по уровню внутризаводского кооперирования (рис. 1). Далее кратко опишем каждый шаг процедур.

Шаг 1. На первом этапе осуществляется постановка задачи классификационного анализа с учетом исходной базы данных x ($x = [1, p]$) по объектам — видам производств a ($a = [1, r]$):

x_1 — доля расходов на вспомогательные производственные операции в общих расходах на производственный процесс, %;

x_2 — расходы на аренду производственного оборудования в расчете на единицу произведенной продукции, рублей;

x_3 — стоимость работ ремонтного хозяйства в расчете на единицу произведенной продукции, рублей;

x_4 — выпуск продукции с единицы ресурсов на реализацию производственного процесса, рублей.

Исследуемый массив данных или объекты классификации — предприятия по производству химической продукции, нефтепродуктов, пластмассовых и резинотехнических изделий и их подвиды ($a = [1, r]$, где $r = 26$).

Шаг 2. Осуществляется верификация данных путем их статистической обработки и выбор переменных: категориальной зависимой переменной (Pc_dep), категориальных (Pc) и непрерывных (Pu) предикторов.

Установим следующее условие для перехода базового показателя ($x_1...x_n$) в категорию переменной:

при отклонении дисперсии выборки (Sv) от среднего значения (Av) более чем в 10 раз базовый показатель x_i не может участвовать в классификационном анализе «Случайный лес»:

$$S_y = \sum \frac{(x_i - x_{av}) \cdot 2}{(n - 1)}, \quad (1)$$

$$\text{если } Sv_x / Av_x > 10, \text{ то } x \notin Pc_{dep}, Pc, Pu, \quad (2)$$

где x_{av} — выборочное среднее значение по показателю;

x_i — i -й элемент выборочной совокупности по показателю;

n — размер выборочной совокупности по показателю.

Шаг 3. Реализация алгоритма «Случайный лес» ансамблевым методом бутстрэп-агрегирования включает пять последовательных процедур.

3.1. Определение основных параметров классификации объектов: количество деревьев (t), количество заданных параметров для выбора расщепления (n_{ss}), максимальная глубина деревьев (max_td), критерий расщепления (C_j).

В качестве критерия расщепления вершины дерева используется критерий Джини (G_i):

$$G_i = 1 - \sum_{j=1}^v P^2(Y_h), \quad (3)$$

где $P(Y_h)$ — удельный вес объектов класса Y_h в подвыборке вершины дерева t , $h = [1, v]$.

3.2. Для каждого дерева (t) из обучающей выборки генерируется подвыборка Z_t , содержащая S_t объектов. Формирование подвыборки Z_t осуществляется на основе случайного выбора с возможным повторением объектов.

3.3. Производится расщепление построенных деревьев t . В соответствии с формулой (3) при бинарной классификации показатель качества расщепления оценивается следующим образом:

$$G_i^{split} = \frac{N_1}{N} G_{t1} + \frac{N_2}{N} G_{t2} \rightarrow \min, \quad (4)$$

где N — число объектов в текущей вершине дерева t (вершина «родитель»);

N_1, N_2 — число объектов в вершинах t_1 и t_2 , соответствующих левому и правому вершинам (вершина «дочь») в случае бинарного дерева.

3.4. На заключительном этапе проводится построение дерева (t) до исчерпания подвыборки Z_t , т.е. до единственного представителя в вершине дерева.

3.5. Итоговый классификатор «Случайный лес» $a(Z_{it})$ выбирает решение по большинству голосов построенных решающих деревьев:

$$a(Z_{it}) = \text{sign} \sum_{j=1}^r b(Z_{it}), \quad (5)$$

где $a(Z_{it})$ — решение итогового классификатора j -го дерева t ($j = 1, t$);

$b(Z_{it})$ — решение базового классификатора j -го дерева ($j = 1, t$);

sign — функция, возвращающая знак своего аргумента.

Шаг 4. Производится оценка качества алгоритма: коэффициент ошибочной классификации (K_m), оценка риска для обучающей и тестовой выборок (A_r):

$$A_r = 1 - \frac{P_{rs}}{P_s}, \quad (6)$$

где A_r — оценка риска ошибки классификации объектов;

P_{rs} — число случаев, правильно классифицированных по деревьям;

P_s — общее число случаев классификации объектов (размер выборки).

Шаг 5. На заключительном этапе производится вывод результатов классификации производств по уровню внутривзаводского кооперирования:

5.1. Построение матрицы распределения производств при количестве решающих деревьев с наименьшим риском ошибочной классификации.

5.2. Оценка полезности моделей на основе диаграмм кумулятивного подъема.

5.3. Вывод итоговых данных о классах производств по наиболее информативным предикторам по критерию G_i .

На этом задача классификационного анализа считается завершенной.

Таблица 1

Дескриптивная статистика данных для классификационного анализа производств по уровню внутривзаводского кооперирования

	x_1	x_2	x_3	x_4
Дисперсия выборки (Sv)	24,820	0,0016	0,0001	0,042
Стандартная ошибка (Es)	0,977	0,0079	0,0015	0,040
Стандартное отклонение (Ds)	4,982	0,0401	0,0076	0,205
Среднее (Av)	5,324	0,0191	0,0034	1,041
Экссесс (Ex)	5,584	14,9134	8,3793	0,402
Асимметричность (As)	2,148	3,6810	2,8012	-0,935
Интервал (Int)	22,960	0,1926	0,0324	0,774
Минимум (Min)	0,000	0,0000	0,0000	0,567
Максимум (Max)	22,960	0,1926	0,0324	1,341
Количество объектов (Ra)	26	26	26	26

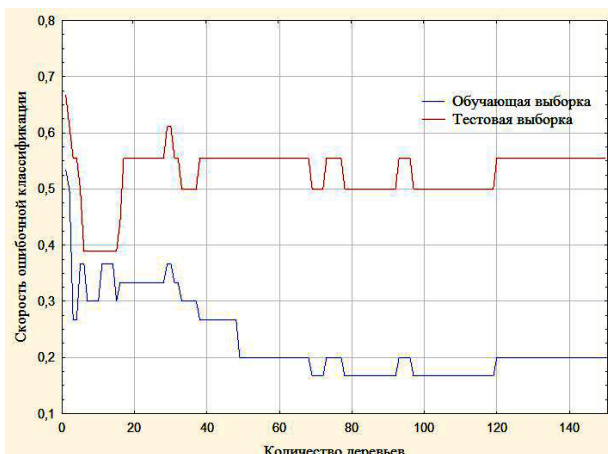


Рис. 2. График коэффициента ошибочной классификации при максимальном количестве решающих деревьев $t_{max} = 150$

Результаты апробации методики. Стоимость организационно-технического обслуживания основного производства можно оценить как отношение затрат на вспомогательные производственные

операции к общим расходам на производственный процесс. Задача рациональной организации вспомогательных процессов на предприятии заключается в снижении их себестоимости до оптимального значения.

Для классификационного анализа нами принята равная для всех категорий цена ошибочной классификации, т.е. матрица стоимости неправильной классификации будет симметричной. Априорное распределение вероятностей величины примем как вероятность того, что объект попадет в один из классов.

Для верификации данных проведен дескриптивный анализ показателей выборки. Статистическая обработка данных показала, что все показатели выборки соответствуют условиям для входных данных ансамблевого метода машинного обучения «Случайный лес» (табл. 1). В качестве зависимой категориальной переменной Pc_dep принят уровень организации основных и вспомогательных производств (текстовая переменная). Независимые категориальные и непрерывные предикторы следующие: $Pc1$ — уровень развития технического аутсорсинга производства (текстовая переменная); $Pu2$ — доля расходов на вспомогательные производственные операции в общих расходах на про-

Таблица 2

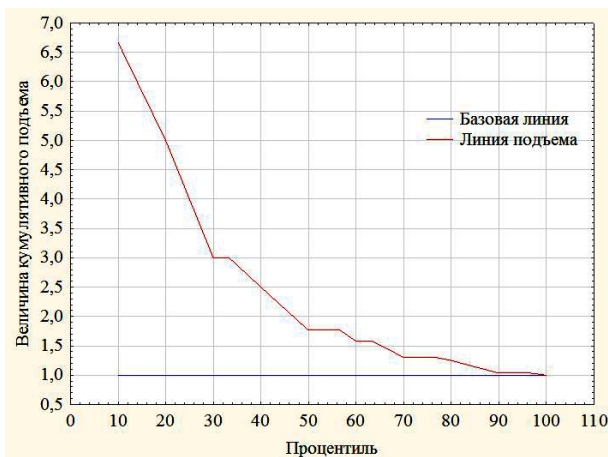
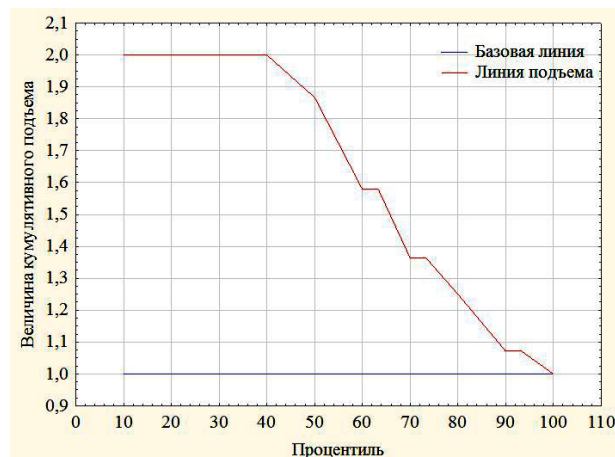
Оценка рисков для обучающей и тестовой выборки при построении случайного леса с разным количеством деревьев

Число деревьев в «Случайном лесу» (t_{max})	Наименование выборки	Оценка риска (A_r)	Стандартная ошибка (E_s)
50	Обучающая выборка	0,083698	0,024870
	Тестовая выборка	0,308521	0,047390
100	Обучающая выборка	0,120413	0,024146
	Тестовая выборка	0,285767	0,047390
150	Обучающая выборка	0,079739	0,020099
	Тестовая выборка	0,305897	0,041889

Таблица 3

Смоделированные классы производств по уровню внутризаводского кооперирования производственных процессов методом «Случайный лес»

Наименование решающих переменных	Тип 1 «Высокий уровень внутризаводского кооперирования, масштабный технический аутсорсинг, высокая техническая эффективность производства»	Тип 2 «Средний уровень внутризаводского кооперирования, частичный технический аутсорсинг, средняя техническая эффективность производства»	Тип 3 «Низкий уровень внутризаводского кооперирования, переход на технический аутсорсинг, низкая техническая эффективность производства»
Pc_dep	$Pc_dep < 5,0$	$5,0 < Pc_dep < 15,0$	$Pc_dep > 15,0$
$Pu3$	0,024	0,011	0,008
$Pu4$	0,0	0,0037	0,0035
$Pu5$	1,238	1,102	1,001
Распределение химических производств	производства: красок и лаков, агрохимических продуктов, изделий из вулканизированной резины и пластмассы, резиновых шин, фотопластинок и фотопленок, парфюмерных и косметических средств, синтетических смол, химических волокон	производства: химических продуктов, пластмассовых изделий для строительства, органических поверхностно-активных веществ, лекарственных препаратов и фармацевтических субстанций, промышленных газов, смазочных материалов, мыла и моющих средств	производство удобрений и азотных соединений

а) Уровень переменной Pc_dep «Тип 1»б) Уровень переменной Pc_dep «Тип 2»Рис. 3. Диаграммы кумулятивного подъема для оценки полезности модели «Случайный лес» по уровням категориальной зависимой переменной (Pc_dep)

изводственный процесс; $Pu3$ — расходы на аренду производственного оборудования в расчете на единицу произведенной продукции; $Pu4$ — стоимость работ ремонтного хозяйства в расчете на единицу произведенной продукции; $Pu5$ — техническая эффективность производства (выпуск продукции с единицы использованных ресурсов).

Далее представлены результаты оценки качества алгоритма классификационного анализа (рис. 2), где показаны коэффициенты ошибочной классификации (Km) по последовательным шагам добавления деревьев для данных обучения и данных тестирования.

В тестовых выборках при вариантах исследуемых множеств решающих деревьев ($t_{max} = 50, 100, 150$) практически в 30 % случаев существует риск неверной классификации деревьев (табл. 2).

Относительно высокое значение риска и существенное отличие вероятности ошибки алгоритма тестовой и обучающей выборки свидетельствует, вероятно, о переобучении алгоритма. Исследователи в области машинного обучения обозначают, что выборки с большим шумом или заданным массивом данных делают модель «Случайный лес» склонной к переобучению [12]. В нашем случае причиной переобучения является высокая сложность модели ввиду стохастической зависимости между предикторами и ответом. В качестве оптимального значения деревьев принимаем $t = 100$ ввиду наименьшей оценки риска ошибочной классификации.

В результате выполнения всех процедур заданная выборка производств классифицирована на три группы (табл. 3):

— тип 1 «Высокий уровень внутризаводского кооперирования, масштабный технический аутсорсинг, высокая техническая эффективность производства»;

— тип 2 «Средний уровень внутризаводского кооперирования, частичный технический аутсорсинг, средняя техническая эффективность производства»;

— тип 3 «Низкий уровень внутризаводского кооперирования, переход на технический аутсорсинг, низкая техническая эффективность производства».

Для подтверждения результатов классификации построены диаграммы кумулятивного подъема по первому и второму типу предприятий, а для

«Типа 3» не выбран ни один случай положительных наблюдений (рис. 3). Наибольший подъем кривой наблюдается при уровнях переменной Pc_dep «Тип 1», где значения оси Y на первых 10 процентах достигают 6,5. Однако данные модели показывают существенный спад кривой подъема после 20–30 % и меньшую вероятность классификации. В графике при уровне переменной Pc_dep «Тип 2» угол линии подъема более приближен к 45°. Соответственно, область между линией подъема и базовой линией наибольшая, что характеризует данную модель как наиболее продуктивную с максимальной вероятностью правильной классификации.

Таким образом, тип 1 включает 17 видов производств и характеризуется долей расходов на вспомогательные производственные операции в общих расходах на производственный процесс менее 5,0 минимальным уровнем технической независимости производства или развитием аутсорсинга технического обслуживания ($Pu3 = 0,024$), отсутствием собственных ремонтных работ, наибольшей технической эффективностью производства $Pu5 = 1,238$.

К типу 2 со средним уровнем внутризаводского кооперирования отнесены 8 видов производств с долей расходов на вспомогательные производственные операции от 5,0 до 15,0 уровнем технической независимости производства $Pu3 = 0,011$ средней стоимостью работ ремонтного хозяйства в расчете на единицу произведенной продукции $Pu4 = 0,0037$ технической эффективностью производства $Pu5 = 1,102$.

Низкий уровень внутризаводского кооперирования основных и вспомогательных производственных процессов (тип 3) наблюдается только в производстве удобрений и азотных соединений: $Pc_dep > 15,0$, $Pu3 = 0,008$, $Pu4 = 0,0035$, $Pu5 = 1,001$.

Заключение. Таким образом, в процессе решения задачи классификации производственных процессов по уровню внутризаводского кооперирования методом алгоритма машинного обучения получены следующие результаты:

1. Предложена авторская методика последовательных действий проведения классификационного анализа для целей выделения объектов с различным уровнем внутризаводского кооперирования основных и вспомогательных производственных

процессов. Методика основана на методе деревьев классификации «Случайный лес» с использованием метаалгоритма композиционного обучения машин Бэггинга. Выбор метода обоснован его наилучшим поведением при большом числе предикторных переменных. Для удобства потенциальных пользователей методика представлена в виде алгоритма из пяти ключевых процедур.

2. Продемонстрирована апробация алгоритма классификационного анализа на примере типологии химических производств по характеру организации вспомогательного хозяйства. Сформирована база данных параметров по 26 видам химических производств, описывающая затраты на вспомогательные операции, расходы на ремонтное хозяйство и обслуживание оборудования, уровень технической эффективности производства. Показана процедура статистической обработки данных на основе дескриптивной аналитики. Проведена оценка качества алгоритма классификационного анализа на основе коэффициентов ошибочной классификации. Установлено оптимальное значение деревьев $t = 100$ ввиду наименьшей оценки риска ошибочной классификации.

3. В результате выполнения всех процедур заданная выборка производств классифицирована на три типа по наиболее информативным переменным. Образованные типы производств характеризуются определенными стандартами значений параметров организации основных и вспомогательных производственных процессов: уровнем внутривозвратской кооперации, развития технического аутсорсинга, технической эффективности производства. Для оценки качества моделей построены диаграммы кумулятивного подъема, где на основе логистической регрессии с максимальной вероятностью правильной классификации определен тип производства 2 «Средний уровень внутривозвратской кооперирования, частичный технический аутсорсинг, средняя техническая эффективность производства».

Полученные данные являются первичной информацией для оценки эффективности организации вспомогательного хозяйства на предприятии, принятия решений о проведении реинжиниринга вспомогательных процессов с целью усиления внутривозвратского кооперирования и снижения уровня затрат.

Библиографический список

1. Травин Г. М., Травин М. М. Системы управления вспомогательным производством промышленных предприятий // Технологии и качество. 2018. № 3 (41). С. 31–35. EDN: YSWYPR.
2. Пигилова Р. Н., Малышева Т. В. Вспомогательное производство: проектирование и эксплуатация систем энергообеспечения // Компетентность. 2022. № 9-10. С. 44–48. DOI: 10.24412/1993-8780-2022-9-44-48. EDN: VBPAUV.
3. Яковлев М. М. Горизонтальная и вертикальная интеграция процессов производства на предприятии // Научный аспект. 2021. Т. 1, № 4. С. 88–92. EDN: FOCHEC.
4. Ключков Ю. С., Конников Е. А. Трансформация структуры производственной себестоимости на промышленном предприятии в процессе интеграции аддитивных технологий // Известия Самарского научного центра Российской академии наук. 2019. Т. 21, № 1 (87). С. 5–10. EDN: WNXXCI.
5. Zhang C., Wang W., Liu L. [et al.]. Three-Branch Random Forest Intrusion Detection Model // Mathematics. 2022. Vol. 10 (23). P. 4460. DOI: 10.3390/math10234460.
6. Клячкин В. Н., Карпунина И. Н. Особенности диагностики технических систем с использованием мультимасштабной классификации // Надежность и качество сложных систем. 2022. № 2 (38). С. 45–52. DOI: 10.21685/2307-4205-2022-2-5. EDN: YXIVPA.
7. Zhao L., Zhu Y., Zhao T. Deep Learning-Based Remaining Useful Life Prediction Method with Transformer Module and Random Forest // Mathematics. 2022. Vol. 10 (16). P. 2921. DOI: 10.3390/math10162921.
8. Пашенко А. Ф., Тордия М. Д., Хижинская Л. Д. Обзор современных подходов к задачам выбора информативных признаков // Датчики и системы. 2020. № 12 (253). С. 16–26. DOI: 10.25728/datsys.2020.12.2. EDN: FFRRGF.
9. Ревина И. В., Бояркин Г. Н. Имитационное моделирование производственного процесса изготовления деталей // Омский научный вестник. 2018. № 6 (162). С. 230–234. DOI: 10.25206/1813-8225-2018-162-230-234. EDN: YSWONV.
10. Лоскутов И. А., Искандарова В. Г. Экосистемный контроллинг как решение сложной задачи оптимизации технологического процесса промышленного производства // Омский научный вестник. 2023. № 1 (185). С. 53–57. DOI: 10.25206/1813-8225-2023-185-53-57. EDN: KEAOAC.
11. Гайфуллина М. М., Низамова Г. З., Гареева З. А. Формирование эффективной системы контроллинга при реструктуризации вспомогательного производства нефтехимического предприятия // Электронный научный журнал Нефтегазовое дело. 2018. № 5. С. 109–133. DOI: 10.17122/ogbus-2018-5-109-133. EDN: VKKUZL.
12. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. 2nd ed. Springer, 2009. 764 p. ISBN 978-0-387-84857-0.

МАЛЫШЕВА Татьяна Витальевна, доктор технических наук, доцент (Россия), профессор кафедры логистики и управления Казанского национального исследовательского технологического университета, г. Казань.
SPIN-код: 6114-7357
AuthorID (РИНЦ): 769164
ORCID: 0000-0003-4306-0188
AuthorID (SCOPUS): 57190414555
ResearcherID: AAM-2396-2021
Адрес для переписки: tv_malysheva@mail.ru

Для цитирования

Малышева Т. В. Разработка алгоритма классификации производств по типу внутривозвратского кооперирования основных и вспомогательных процессов методом машинного обучения // Омский научный вестник. 2024. № 1 (189). С. 12–19. DOI: 10.25206/1813-8225-2024-189-12-19.

Статья поступила в редакцию 10.07.2023 г.

© Т. В. Малышева

DEVELOPMENT OF THE ALGORITHM FOR CLASSIFYING INDUSTRIES ACCORDING TO THE TYPE OF INTRA-FACTORY COOPERATION OF MAIN AND AUXILIARY PROCESSES USING MACHINE LEARNING

The task of rational organization of auxiliary processes at the enterprise is to reduce their cost by deep integration into the main production process. The purpose of the article is to develop a classification analysis algorithm for assessing the dependencies between the main and auxiliary units and the typology of production processes according to the level of intra-factory cooperation. As a method for determining the type of production, the Random Forest machine learning method using the bagging machine learning meta-algorithm is proposed. Parameters have been developed that describe the costs of auxiliary operations, the costs of repair facilities and equipment maintenance, the level of technical efficiency of production. Approbation of the algorithm on the example of chemical enterprises made it possible to distinguish three types of production according to the nature of intra-plant cooperation of processes according to the most informative parameters. To assess the usefulness and performance of the models, cumulative lift diagrams are constructed, where the most productive type is determined with an average level of intra-factory cooperation. The results are the primary diagnostics of the organization of auxiliary facilities, decision-making on the reengineering of processes in order to strengthen intra-factory cooperation and reduce costs.

Keywords: auxiliary production, production processes, intra-factory cooperation, algorithm, classification analysis, machine learning, random forest.

References

1. Travin G. M., Travin M. M. Sistemy upravleniya vspomogatel'nym proizvodstvom promyshlennykh predpriyatii [Production auxiliary process control systems] // *Tekhnologii i kachestvo. Technologies & Quality*. 2018. No. 3 (41). P. 31–35. EDN: YSWYPR. (In Russ.).
2. Pigilova R. N., Malysheva T. V. Vspomogatel'noye proizvodstvo: proyektirovaniye i ekspluatatsiya sistem energoobespecheniya [Auxiliary production: design and operation of energy supply systems] // *Kompetentnost'. Journal Kompetentnost'*. 2022. No. 9–10. P. 44–48. DOI: 10.24412/1993-8780-2022-9-44-48. EDN: VBPAYV. (In Russ.).
3. Yakovlev M. M. Gorizontal'naya i vertikal'naya integratsiya protsessov proizvodstva na predpriyatii [Horizontal and vertical integration of production processes at the enterprise] // *Nauchnyy aspekt. Nauchnyy Aspekt*. 2021. Vol. 1, no. 4. P. 88–92. EDN: FOCHES. (In Russ.).
4. Klochkov Yu. S., Konnikov E. A. Transformatsiya struktury proizvodstvennoy sebestoimosti na promyshlennom predpriyatii v protsesse integratsii additivnykh tekhnologiy [Changes in the structure of production costs at an industrial enterprise while integrating additive technologies] // *Izvestiya Samarskogo Nauchnogo Tsentra Rossiyskoy Akademii Nauk. Izvestia of Samara Scientific Center of the Russian Academy of Sciences*. 2019. Vol. 21, no. 1 (87). P. 5–10. EDN: WNXXCI. (In Russ.).
5. Zhang C., Wang W., Liu L. [et al.]. Three-Branch Random Forest Intrusion Detection Model // *Mathematics*. 2022. Vol. 10 (23). P. 4460. DOI: 10.3390/math10234460. (In Engl.).
6. Klyachkin V. N., Karpunina I. N. Osobennosti diagnostiki tekhnicheskikh sistem s ispol'zovaniem mul'tiklassovoy klassifikatsii [Features of diagnostics of technical systems using multiclass classification] // *Nadezhnost' i Kachestvo Slozhnykh Sistem. Reliability and Quality of Complex Systems*. 2022. No. 2 (38). P. 45–52. DOI: 10.21685/2307-4205-2022-2-5. EDN: YXIVPA. (In Russ.).
7. Zhao L., Zhu Y., Zhao T. Deep Learning-Based Remaining Useful Life Prediction Method with Transformer Module and Random Forest // *Mathematics*. 2022. Vol. 10 (16). P. 2921. DOI: 10.3390/math10162921. (In Engl.).
8. Pashchenko A. F., Tordiya M. D., Khizhinskaya L. D. Obzor sovremennykh podkhodov k zadacham vybora informativnykh priznakov [Review of modern approaches to the problems of informative features selection] // *Datchiki i Sistemy. Sensors & Systems*. 2020. No. 12 (253). P. 16–26. DOI: 10.25728/datsys.2020.12.2. EDN: FFRRGF. (In Russ.).
9. Revina I. V., Boyarkin G. N. Imitatsionnoye modelirovaniye proizvodstvennogo protsessa izgotovleniya detaley [Simulation of production process manufacturing parts] // *Omskiy Nauchnyy Vestnik. Omsk Scientific Bulletin*. 2018. No. 6 (162). P. 230–234. DOI: 10.25206/1813-8225-2018-162-230-234. EDN: YSWONV. (In Russ.).

10. Loskutov I. A., Iskandarova V. G. Ekosistemnyy kontrolling kak resheniye slozhnoy zadachi optimizatsii tekhnologicheskogo protsessa promyshlennogo proizvodstva [Ecosystem controlling as a solution to complex task of optimizing the technological process of industrial production] // Omskiy Nauchnyy Vestnik. *Omsk Scientific Bulletin*. 2023. No. 1 (185). P. 53–57. DOI: 10.25206/1813-8225-2023-185-53-57. EDN: KEAOAC. (In Russ.).

11. Gayfullina M. M., Nizamova G. Z., Gareyeva Z. A. Formirovaniye effektivnoy sistemy kontrollinga pri restrukturizatsii vspomogatel'nogo proizvodstva neftekhimicheskogo predpriyatiya [Formation of effective system of controlling when restructuring auxiliary production of the petrochemical enterprise] // Elektronnyy Nauchnyy Zhurnal Neftegazovoye Delo. *Electronic Scientific Journal Oil and Gas Business*. 2018. No. 5. P. 109–133. DOI: 10.17122/ogbus-2018-5-109-133. EDN: VKKUZI. (In Russ.).

12. Hastie T., Tibshirani R., Friedman J. *The Elements of Statistical Learning*. 2nd ed. Springer, 2009. 764 p. ISBN 978-0-387-84857-0. (In Engl.).

MALYSHEVA Tatyana Vitalievna, Doctor of Technical Sciences, Associate Professor, Professor of Logistics and

Management Department, Kazan National Research Technological University, Kazan.

SPIN-code: 6114-7357

AuthorID (RSCI): 769164

ORCID: 0000-0003-4306-0188

AuthorID (SCOPUS): 57190414555

ResearcherID: AAM-2396-2021

Correspondence address: tv_malysheva@mail.ru

For citations

Malysheva T. V. Development of the algorithm for classifying industries according to the type of intra-factory cooperation of main and auxiliary processes using machine learning // *Omsk Scientific Bulletin*. 2024. No. 1 (189). P. 12–19. DOI: 10.25206/1813-8225-2024-189-12-19.

Received July 10, 2023.

© T. V. Malysheva